SEQUENCE LISTING

<110> Merck & Co., Inc.

<120> OPTIMIZED EXPRESSION OF HPV 52 L1 IN
   YEAST

<130> 21571 PCT

<150> 60/555926
<151> 2004-03-24

<160> 7

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1512
<212> DNA
<213> Artificial Sequence

<220>
<223> HPV52L1R

<400> 1
atgtccgtct ggagaccatc cgaagctact gtctacttgc caccagttcc agtctctaag 60
gttgtctcta ccgacgaata cgtctccaga acctccatct actactacgc tggttcctct 120
agattgttga ctgtcggtca cccatacttc tctatcaaga acacctcctc cggtaacggt 180
aagaaggtct tggttccaaa ggtctctggt ttgcaataca gagtcttcag aatcaagttg 240
ccagacccaa acaagttcgg tttcccagac actagtttct acaacccaga aactcaaaga 300
ttggtctggg cttgtactgg tttggaaatc ggtagaggtc aaccattggg tgtcggtatc 360
tctggtcacc cattgttgaa caagttcgac gacactgaaa cctctaacaa gtacgctggt 420
aagccaggta tcgataacag agaatgtttg tctatggact acaagcaaac tcaattgtgt 480
atcttgggtt gtaagccacc aatcggtgaa cactggggta agggtactcc atgtaacaac 540
aactctggta acccaggtga ctgtccacca ttgcaattga tcaactccgt catccaagac 600
ggtgacatgg tcgacactgg tttcggttgt atggacttca acaccttgca agcttctaag 660
tccgacgtcc caatcgacat ctgttcctct gtctgtaagt acccagacta cttgcaaatg 720
gcttctgaac atacggtga ctccttgttc ttcttcttga agagaaca aatgttcgtc 780
agacacttct tcaacagagc tggtaccttg ggtgacccag ttccaggtga cttgtacatc 840
caaggttcca actctggtaa cactgctact gtccaatcct ctgctttctt cccaactcca 900
tctggttcca tggtcacctc cgaatcccaa ttgttcaaca agccatactg gttgcaaaga 960
gctcaaggtc acaacaacgg tatctgttgg ggtaaccaat gttcgtcac cgtcgtcgac 1020
actactagat ctactaacat gaccttgtgt gctgaagtca agaaggaatc cacctacaag 1080
aacgaaaact caaggaata cttgagacac ggtgaagaat cgacttgca attcatcttc 1140
caattgtgta agatcacctt gaccgctgac gtcatgactt acatccacaa gatggacgct 1200
actatcttgg aagactggca attcggtttg actccaccac catccgcttc cttggaagac 1260
acttacagat cgtcacttc cactgctatc acctgtcaaa gaacactcc accaaaggt 1320
aaggaagacc cattgaagga ctacatgttc tgggaagtcg acttgaagga aaagttctct 1380
gctgacttgg accaattccc attgggtaga aagttcttgt gcaagctgg tttgcaagct 1440
agaccaaagt tgaagagacc agctagctct gctccaagaa cttccaccaa gaagaagaag 1500
gtcaagagat aa   1512

<210> 2
<211> 503

```
<212> PRT
<213> Human papillomavirus type 52

<400> 2
Met Ser Val Trp Arg Pro Ser Glu Ala Thr Val Tyr Leu Pro Pro Val
1               5                   10                  15
Pro Gly Ser Lys Val Val Ser Thr Asp Glu Tyr Val Ser Arg Thr Ser
            20                  25                  30
Ile Tyr Tyr Tyr Ala Gly Ser Ser Arg Leu Leu Thr Val Gly His Pro
        35                  40                  45
Tyr Phe Ser Ile Lys Asn Thr Ser Ser Gly Asn Gly Lys Lys Val Leu
    50                  55                  60
Val Pro Lys Val Ser Gly Leu Gln Tyr Arg Val Phe Arg Ile Lys Leu
65                  70                  75                  80
Pro Asp Pro Asn Lys Phe Gly Phe Pro Asp Thr Ser Phe Tyr Asn Pro
                85                  90                  95
Glu Thr Gln Arg Leu Val Trp Ala Cys Thr Gly Leu Glu Ile Gly Arg
            100                 105                 110
Gly Gln Pro Leu Gly Val Gly Ile Ser Gly His Pro Leu Leu Asn Lys
            115                 120                 125
Phe Asp Asp Thr Glu Thr Ser Asn Lys Tyr Ala Gly Lys Pro Gly Ile
    130                 135                 140
Asp Asn Arg Glu Cys Leu Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys
145                 150                 155                 160
Ile Leu Gly Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Thr
                165                 170                 175
Pro Cys Asn Asn Asn Ser Gly Asn Pro Gly Asp Cys Pro Pro Leu Gln
            180                 185                 190
Leu Ile Asn Ser Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe
            195                 200                 205
Gly Cys Met Asp Phe Asn Thr Leu Gln Ala Ser Lys Ser Asp Val Pro
    210                 215                 220
Ile Asp Ile Cys Ser Ser Val Cys Lys Tyr Pro Asp Tyr Leu Gln Met
225                 230                 235                 240
Ala Ser Glu Pro Tyr Gly Asp Ser Leu Phe Phe Phe Leu Arg Arg Glu
                245                 250                 255
Gln Met Phe Val Arg His Phe Phe Asn Arg Ala Gly Thr Leu Gly Asp
            260                 265                 270
Pro Val Pro Gly Asp Leu Tyr Ile Gln Gly Ser Asn Ser Gly Asn Thr
            275                 280                 285
Ala Thr Val Gln Ser Ser Ala Phe Phe Pro Thr Pro Ser Gly Ser Met
    290                 295                 300
Val Thr Ser Glu Ser Gln Leu Phe Asn Lys Pro Tyr Trp Leu Gln Arg
305                 310                 315                 320
Ala Gln Gly His Asn Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val
                325                 330                 335
Thr Val Val Asp Thr Thr Arg Ser Thr Asn Met Thr Leu Cys Ala Glu
            340                 345                 350
Val Lys Lys Glu Ser Thr Tyr Lys Asn Glu Asn Phe Lys Glu Tyr Leu
            355                 360                 365
Arg His Gly Glu Glu Phe Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys
    370                 375                 380
Ile Thr Leu Thr Ala Asp Val Met Thr Tyr Ile His Lys Met Asp Ala
385                 390                 395                 400
```

```
Thr Ile Leu Glu Asp Trp Gln Phe Gly Leu Thr Pro Pro Pro Ser Ala
            405                     410                 415
Ser Leu Glu Asp Thr Tyr Arg Phe Val Thr Ser Thr Ala Ile Thr Cys
            420                     425                 430
Gln Lys Asn Thr Pro Pro Lys Gly Lys Glu Asp Pro Leu Lys Asp Tyr
            435                     440                 445
Met Phe Trp Glu Val Asp Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp
    450                     455                 460
Gln Phe Pro Leu Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Gln Ala
465                     470                 475                 480
Arg Pro Lys Leu Lys Arg Pro Ala Ser Ser Ala Pro Arg Thr Ser Thr
                485                     490                 495
Lys Lys Lys Lys Val Lys Arg
                500
```

```
<210> 3
<211> 1512
<212> DNA
<213> Human Papillomavirus Type 52
```

```
<400> 3
atgtccgtgt ggcggcctag tgaggccact gtgtacctgc ctcctgtacc tgtctctaag 60
gttgtaagca ctgatgagta tgtgtctcgc acaagcatct attattatgc aggcagttct 120
cgattactaa cagtaggaca tccctatttt tctattaaaa acaccagtag tggtaatggt 180
aaaaaagttt tagttcccaa ggtgtctggc ctgcaataca gggtatttag aattaaattg 240
ccggacccta ataaatttgg ttttccagat acatcttttt ataacccaga aacccaaagg 300
ttggtgtggg cctgtacagg cttggaaatt ggtaggggac agcctttagg tgtgggtatt 360
agtgggcatc ctttattaaa caagtttgat gatactgaaa ccagtaacaa atatgctggt 420
aaacctggta tagataatag ggaatgttta tctatggatt ataagcagac tcagttatgc 480
attttaggat gcaaacctcc tataggtgaa cattggggta agggaacccc ttgtaataat 540
aattcaggaa atcctgggga ttgtcctccc ctacagctca ttaacagtgt aatacaggat 600
ggggacatgg tagatacagg atttggttgc atggatttta ataccttgca agctagtaaa 660
agtgatgtgc ccattgatat atgtagcagt gtatgtaagt atccagatta tttgcaaatg 720
gctagcgagc catatggtga cagtttgttc ttttttctta gacgtgagca aatgtttgtt 780
agacactttt ttaatagggc cggtacctta ggtgaccctg tgccaggtga tttatatata 840
caagggtcta actctggcaa tactgccact gtacaaagca gtgctttttt tcctactcct 900
agtggttcta tggtaacctc agaatcccaa ttatttaata aaccgtactg gttacaacgt 960
gcgcagggcc acaataatgg catatgttgg ggcaatcagt tgtttgtcac agttgtggat 102 0
accactcgta gcactaacat gactttatgt gctgaggtta aaaggaaag cacatataaa 108 0
aatgaaaatt ttaaggaata ccttcgtcat ggcgaggaat ttgatttaca atttattttt 114 0
caattgtgca aaattacatt aacagctgat gttatgacat acattcataa gatggatgcc 120 0
actattttag aggactggca atttggcctt accccaccac cgtctgcatc tttggaggac 126 0
acatacagat ttgtcacttc tactgctata acttgtcaaa aaacacgcc acctaaagga 132 0
aaggaagatc ctttaaagga ctatatgttt tgggaggtgg atttaaaaga aaagttttct 138 0
gcagatttag atcagtttcc tttaggtagg aagttttttgt tacaggcagg gctacaggct 144 0
aggcccaaac taaaacgccc tgcatcatcg gccccacgta cctccacaaa gaagaaaaag 150 0
gttaaaaggt aa                                                         151 2
```

```
<210> 4
<211> 21
<212> DNA
<213> Artificial Sequence
```

```
<220>
<223> PCR Primer

<400> 4
atgtccgtgt ggcggcctag t                                                           21


<210> 5
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 5
gagatctcaa ttacacaaag tg                                                          22


<210> 6
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 6
gagatctcac aaaacaaaat gtccgtgtgg c                                                 31


<210> 7
<211> 1512
<212> DNA
<213> Artificial Sequence

<220>
<223> 52L1R antisense

<400> 7
tacaggcaga cctctggtag gcttcgatga cagatgaacg gtggtcaagg tcagagattc 60
caacagagat ggctgcttat gcagaggtct tggaggtaga tgatgatgcg accaaggaga 120
tctaacaact gacagccagt gggtatgaag agatagttct tgtggaggag gccattgcca 180
ttcttccaga accaaggttt ccagagacca aacgttatgt ctcagaagtc ttagttcaac 240
ggtctgggtt tgttcaagcc aaagggtctg tgatcaaaga tgttgggtct ttgagtttct 300
aaccagaccc gaacatgacc aaacctttag ccatctccag ttggtaaccc acagccatag 360
agaccagtgg gtaacaactt gttcaagctg ctgtgacttt ggagattgtt catgcgacca 420
ttcggtccat agctattgtc tcttacaaac agatacctga tgttcgtttg agttaacaca 480
tagaacccaa cattcggtgg ttagccactt gtgaccccat tcccatgagg tacattgttg 540
ttgagaccat tgggtccact gacaggtggt aacgttaact agttgaggca gtaggttctg 600
ccactgtacc agctgtgacc aaagccaaca tacctgaagt gtggaacgt tcgaagattc 660
aggctgcagg gttagctgta gacaaggaga cagacattca tgggtctgat gaacgtttac 720
cgaagacttg gtatgccact gaggaacaag aagaagaact cttctcttgt ttacaagcag 780
tctgtgaaga agttgtctcg accatggaac ccactgggtc aaggtccact gaacatgtag 840
gttccaaggt tgagaccatt gtgacgatga caggttagga gacgaaagaa gggttgaggt 900
agaccaaggt accagtggag gcttagggtt aacaagttgt tcggtatgac caacgtttct 960
cgagttccag tgttgttgcc atagacaacc ccattggtta acaagcagtg cagcagctg 1020
tgatgatcta gatgattgta ctggaacaca cgacttcagt tcttccttag gtggatgttc 1080
```

                                               - 4 -

```
ttgcttttga agttccttat gaactctgtg ccacttctta agctgaacgt taagtagaag 1140
gttaacacat tctagtggaa ctggcgactg cagtactgaa tgtaggtgtt ctacctgcga 1200
tgatagaacc ttctgaccgt taagccaaac tgaggtggtg gtaggcgaag gaaccttctg 1260
tgaatgtcta agcagtgaag gtgacgatag tggacagttt tcttgtgagg tggtttccca 1320
ttccttctgg gtaacttcct gatgtacaag acccttcagc tgaacttcct tttcaagaga 1380
cgactgaacc tggttaaggg taacccatct ttcaagaaca acgttcgacc aaacgttcga 1440
tctggtttca acttctctgg tcgatcgaga cgaggttctt gaaggtggtt cttcttcttc 1500
cagttctcta tt                                              1512
```